Knowledge Based Systems in DMA's Digital Production System

Maria Gruenwald

DMA Systems Center
8613 Lee Highway
Fairfax, VA 22031

N/A

DMASC/MGE

N/A

AD-A230 771

To be presented at ACSM/ASPRS Auto Carto Conference, Baltimore, MD, 25-29 March 1991

Publicly Available
a.) Paper to be published in findings
b.) Presnetation to be delivered at conference

DTIC
ELECTE
JAN 11 1991
S E D

The Digital Production System (DPS) is an end-to-end system with softcopy source as primary input and Defense Mapping Agency cartographic products as output. Within the DPS, there are Knowledge Based Systems (KBSs) which support the overall process in the planning of production, stereo extraction of cartographic features, and the generation of products.

Digital Production System
Defense Mapping Agency
Knowledge Based Systems

N/A

Unclassified Unclassified Unclassified N/A

Accession For
NTIS GRA&I X
DTIC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

KNOWLEDGE BASE SYSTEMS IN DMA'S DIGITAL PRODUCTION SYSTEM

Maria M. Gruenewald
Jeffrey L. Kretsch, Ph.D.

Defense Mapping Agency
8613 Lee Highway
Fairfax, Virginia 22031

ABSTRACT

The Digital Production System (DPS) is an end-to-end system with softcopy source as primary input and Defense Mapping Agency (DMA) cartographic products as output. Within the DPS, there are Knowledge Based Systems (KBSs) which support the overall process in the planning of production, stereo extraction of cartographic features, and the generation of products.

INTRODUCTION

DMA has made a commitment to the automation of its cartographic processes by the development of the DPS which consists of "segments". One segment, the Production Management Segment, can be thought of as the "brains" of the system. This segment controls and monitors all of the DPS activity. Another segment, the Data Services Segment, acts as the nervous system, carrying messages and data to all parts of the DPS. The Data Services Segment is also responsible for the storage of the DPS Mapping Charting and Geodetic (MC&G) Data Base.

In addition, there are three segments in a production pipeline of activity. The first of these, the Source Preparation Segment, prepares the assignment package which includes instructions, specifications, and the needed source to complete a particular assignment. The second segment in the pipeline, the Data Extraction Segment, takes the sources from the assignment package and extracts all of the terrain and feature data according to the specifications for this assignment. The extracted terrain and feature data are stored in the MC&G Data Base to be retrieved by the third segment in the production pipeline, the Product Generation Segment. The Product Generation Segment takes the assignment package and the MC&G data from the Data Base and actually creates the DMA products specified by this assignment.

DMA AND KNOWLEDGE BASE SYSTEMS

A KBS is sometimes called an "Expert System" because the knowledge in a KBS reflects or emulates the knowledge of an expert in the field of the specific need. In DMA's case, the domain experts are cartographers, air information specialists, bathymetrists, and MC&G production planners. The KBSs in DMA share the employee's tasks, providing a reduction in the intensity of effort in the production and planning activities of the DPS. Also, KBSs, like other computer programs, help to standardize what the cartographer sees which helps to guide what the cartographer does. So we can standardize what many different cartographers do, and contribute to the consistency and quality of DMA products. The KBSs are transparent to use. The KBSs offer suggestions and the users can accept the suggestions or override them.

91 10 08

Designing the KBSs for the DPS was particularly challenging for several reasons, among which were the following:

- There was no single expert.

Usually a KBS models the judgment of an "expert" in the field of the application. While DMA has many employees who are experts at what they do, there is no single "universal expert" whose mind could be tapped for the knowledge to put into the KBSs. The designers had to combine the domain expert knowledge into one "seamless" base of knowledge, modelling the expertise from a variety of different sources that did not always agree on one best way to approach a problem.

- There were no DPS operational experts.

Since DPS was not yet operational, there were no operational experts which could use their knowledge in KBS development. The DMA Product and Extraction Specifications for the DPS had not yet been developed. This was an interesting challenge for the designers who had to work with experts and try to extrapolate their input into what future experts would say. Many DPS concepts and new Specifications not yet in use had to be considered to derive any benefit from the Knowledge Acquisition phase of the KBS development.

Architecturally speaking, the DPS KBSs are embedded within larger applications. The Data Extraction Segment, the Product Generation Segment, and the Production Management Segment of the DPS segments incorporate KBSs to address a portion of their piece of the system.

THE DATA EXTRACTION SEGMENT KNOWLEDGE BASE SYSTEM

The knowledge in the Data Extraction Segment KBS comes from several sources:

- DMA domain experts,
- Data on MC&G feature occurrence frequencies,
- DMA Extraction Specifications

The DMA domain experts were taken from the ranks of cartographers who were experts in their fields. Because of the design challenges mentioned previously, the cartographers who participated in the Knowledge Acquisition task had to first become familiar with the DPS.

The second source of knowledge was the Data on MC&G features occurrence frequencies. In a survey of DMA products, statistics were gathered to determine which features were most frequently occurring across all DMA products. The statistical analysis combined with the DMA domain expert interaction helped to focus the KBS development on the activities that would provide the most assistance during the feature extraction process.

Finally, the DMA Extraction Specifications helped to build knowledge into the KBS about DMA MC&G features, and their attributes, as well as when and how they are to be extracted.

Because of the complexity of MC&G data extraction processes and the need for a specially devised inference engine, a commercially available expert system shell could not be used. The KBS in the Data Extraction Segment uses rules frames and objects to perform Automated Feature Extraction of MC&G features. Automated Feature Extraction also includes interactive processes which facilitate feature extraction.

The Data Extraction cartographer, using the KBS, has the final say as to the outcome of its processing. An override to the KBS can lead to an operator-supplied result or lead back into the KBS for further processing.

The Automated Feature Extraction Function includes:

- Stereo delineation,
- Identification,
- Attribution.

The Stereo Delineation of MC&G Features

This is a semi-automated process accomplished through computer vision and interactive graphic processing. Computer vision is a technology that combines image processing techniques and knowledge about the task to accomplish automated "image understanding". The more knowledge that is incorporated, the greater the degree of understanding that can be achieved. In DMA, the real image understanding is in the mind of the cartographer. Computer vision is aimed at accomplishing feature delineations with less effort on the part of the cartographer.

The computer vision tools are organized as a "tool box" of stereo algorithms. The idea is to provide the right tool for the right feature. The delineations of different types of features are suited to different algorithms. The tool box contains a variety of tools of varying degrees of sophistication and applicability. The varying degrees of sophistication represent varying degrees of image understanding on the part of the KBS. The system was developed over several years. The reliable tools available at the start of the project made way for improved tools. Computer vision became more advanced as the DPS development continued. Insight into this natural evolution of technology convinced the designers of the KBS to employ the "tool box concept". The computer vision tool box is flexible and receptive to new tools as new algorithms became practical to implement. This design approach contributes to the future growth of the system.

The current computer vision tools include:

- Single feature delineation.

Also called "region growing", this tool evaluates a patch of pixels in softcopy source that surround a cartographer-selected point. All pixels in the neighborhood of this patch that have similar grey scales are added to the patch. The patch grows until the neighborhood can no longer be expanded because of new pixel dissimilarity. The success of this tool depends very much on the cartographer's initial point selection.

- Edge follower.

Using a Sobel operator, it works well in high contrast images with well-defined feature edges. When it gets going, it delineates very quickly (about 125 pixels per second).

- Pattern follower.

This tool follows an edge-weighted cross-section of pixels. The pattern is typically between two parallel edges, although this is not a restriction of the algorithm. The pattern follower is a robust tool for delineating roads, railroads, canals, and any other cartographic feature that has a fairly constant width and no sharp inflections. It is robust because it can process through objects that disrupt the pattern, like cars, trucks, and trains. The pattern follower can find its way over bridges, under overpasses, up exit ramps, and across intersections. Like the edge follower, the pattern follower works very fast. It does not limit itself to just the pixels visible on the screen; it continues the delineation to the edge of the image.

- Delineation refinement.

This tool requires as input the delineation of a line or area feature. The input delineation can be very rough, often the result of a previous delineation at a reduced resolution. The vertices of the rough delineation act as control points to the refinement process. This tool refines the delineation of typically large features by applying photometric and geometric filters.

- Supervised classifier.

Also called "region delineation by class", this computer vision tool is based on two neural network software implementations. This tool requires training by the cartographer prior to processing. It works best when the tool is trained to discriminate among as few classes as possible. The tool is retrained each time it is used for each image upon which it is used. It is very good at classifying natural features such as vegetation and water.

Most of the tools in the computer vision tool box operate as background processes. The cartographer provides start-up parameters (which provide more knowledge) to the computer vision tools and the system completes a candidate delineation. The cartographer either accepts the suggested delineation or edits it using the interactive graphic editing tools which process outside of the KBS. The cartographer also has the option to bypass the computer vision tools and use only the interactive tools, but using the computer vision tools promises a time and labor savings in the delineation task. The interactive graphic editing tools are a standard set that includes erase, merge, rubberband, and point-to-point draw.

The Identification of MC&G Features

This part of the Automated Feature Extraction function assists the cartographer in assigning an identification code to a delineated feature. The KBS provides this assistance by applying a series of filters over the list of possible identification codes to narrow down the list of likely choices. Each filter is an application of KBS rules. The filtering activity is a background process and it is never viewed by the cartographer. From the cartographer's perspective, the identification assistance is instantaneous.

One reason for the speed of the identification process is that the KBS is selective about which features it will identify. During the Knowledge Acquisition phase, the feature extraction domain experts identified those features which occur most often across all DMA products. These features are called "priority features" and it is only this subset of all DMA features that are presented to the filters in the identification process. The following are the filters which assist in the feature identification process:

- The metric partition filter.

The metric partition filter operates on the principle that the computer does not "see" features: the computer "sees" sizes, shapes, and heights. The priority features are organized into bins by their expected sizes, shapes, and heights. The delineation of a feature produces size, shape, and height information which is used by this filter to determine the appropriate bin for this delineation. Then, the likely features having that size, shape, and height are filtered from the identification possibilities into the identification choices. The results from each of the filters constrains the list of choices so that the cartographer selects the appropriate identification from a list of eight or less features.

- The extraction specification filter.

This filter limits the list of choices to only those features which are appropriate for the particular assignment, based on the DMA Extraction Specifications for the DPS.

- The context area filter

This filter applies rules about expectations of features types within a particular context (such as Agricultural, Coastal, Air Facility, Excavation-Drilling, Built-Up Area, and Mountainous).

- The geographic region filter.

This filter lists identification choices based on rules about feature expectations in different parts of the world.

Together, these filters reduce the list of reasonable identification choices to eight or fewer feature codes. When presented with the list of likely identification choices, the cartographer may select from the list or choose to enter a different identification, thereby overriding the KBS. These overrides are accumulated automatically by the KBS, and summarized in a Summary Report.

This Summary Report is evaluated by a Knowledge Engineer to determine if there are trends emerging in the KBS performance that require attention. From the information in the Summary Report, the Knowledge Engineer can isolate potential improvements to the KBS, develop candidate changes, and test the viability of those changes. This knowledge engineering activity supports the Data Extraction Segment KBS "learning".

This learning is not automatic. The DPS KBSs does not possess machine learning capabilities, that is, the power to change themselves based on their own evaluation of their performance. In this segment, however, automatic accumulation and summarizing of overrides contributes to the Knowledge Engineer's ability to "learn" where to improve the KBS. As with the image understanding, the real learning is on the part of the human.

The Attribution of MC&G Features

The attribution task begins when the cartographer selects a previously delineated, and identified feature object for attribution. This function uses rules and frames with procedural attachments to enter attributes automatically, to prompt the cartographer for needed values based on what the KBS knows and what the cartographer has told it. This process is dynamic in that it continues to provide support to the cartographer during the input process to automatically infer additional attribute values.

The cartographer is presented with a portion of the feature's object record that represents the attributes for the feature. The KBS makes every effort to enter as many attribute values as it can before and during the cartographer's part in the feature attribution. For this reason, the ordering of attributes and the ordering of prompts to the cartographer are important in the design of the KBS.

The cartographer has the ultimate responsibility for the data that describes the feature. Overrides of attribute suggestions are accumulated into the Summary Report as they are during the identification process.

The Data Extraction Segment KBS assists in the stereo delineation, identification, and attribution of MC&G features. These features are ultimately stored in the MC&G Data Base where they are retrieved and placed on DMA products by the Product Generation Segment.

THE PRODUCT GENERATION SEGMENT KNOWLEDGE BASE SYSTEM

The sources of knowledge in the Product Generation Segment KBS is based on the DMA Product Specifications for the DPS along with the best available cartographic expertise.

The Product Generation Segment uses an object oriented approach with objects, states, message passing, and methods. It makes it difficult to sort out how much of the segment is part of the KBS and how much is an object oriented programming approach that is not knowledge-based. In addition to objects, the design uses rules to facilitate message passing, method computation, and changing object state. The Product Generation Segment required a specially developed inference engine rather than adoption of a commercially available expert system shell for many of the same reasons as with the Data Extraction Segment.

Unlike the Data Extraction Segment KBS which supports only the Automated Feature Extraction function, the Product Generation Segment KBS supports many functions. It provides this support as a distributed service and can be accessed by many different parts of the segment.

Some of the services provided by the KBS in this segment are the following:

- Automatic feature symbolization
- Data segregation
- Automatic feature generalization
- Feature recognition
- Automatic name and label placement
- Automatic conflict detection and resolution

As an example, consider the last service on this list. The Automatic conflict detection and resolution service is invoked when features symbolized on a DMA product are in a coalescent state. Coalescence occurs when features have been symbolized and are ready for placement on a product. It is important that they do not obscure one another, in order to eliminate overprints on the final product, and that their relationships are preserved. This condition is especially problematic when features are represented at a scale that is less detailed than the scale at which they were extracted. The KBS is called upon to detect and resolve the conflict, actually moving the symbolic representations of the features on the product to be within the Specifications.

Although, unlike the Data Extraction Segment, the Product Generation Segment KBS has no override audit capabilities that provide summary information about its performance, the rules within the KBS are kept in special files that are more accessible and easier to fine-tune than conventional software.

While the Data Extraction Segment KBS handles the Automated Feature Extraction task and the Product Generation Segment KBS handles various services in support of the creation of DMA products, the Production Management Segment KBS helps decide which extraction and product generation tasks will be done.

THE PRODUCTION MANAGEMENT SEGMENT KNOWLEDGE BASE SYSTEM

The source of the knowledge in the Production Management Segment KBS is the expertise of DMA production planners. These production planners provided information about how they do their jobs today and how they expect they will do their jobs after the DPS is fully operational.

The Production Management Segment KBS is rule-based. It provides intelligent decision support to the planning of DMA MC&G production activities. These activities include geopositioning, MC&G data extraction, and DMA product finishing. The Production Management Segment uses an expert system shell called "Knowledge Tool" (KTOOL) which is an IBM product.

Before the DPS, production planning was a process that was almost completely manual. Planners had to keep track of all customer requests and all available resources to handle those requests. They also had to know all of the production activities carried out at multiple, geographically distinct production centers that must be accomplished in order to satisfy a customer's request. Combine with this the heuristic that people and equipment should not be idle. Keeping track of all of this manually is a very taxing and labor intensive task for the human production planner.

Add to this complexity the incorporation of Multiple Product Operations (MPO) into the DPS. This consideration in planning allows DMA production to exploit geocoincidence when multiple product requirements overlap in a geographic area. Automating the thought processes that accomplish MPO was a complex problem, but in the DPS, the KBS bears the burden of the complexity instead of the human production planner.

In addition to the creation of a plan for DMA production activities, the KBS in this segment provides management with a useful tool to model different production scenarios. This is a powerful application which can help production planners to project much longer range plans, analyze contingencies, or become better prepared to deal with crises. The "what if" capability allows the production planner to modify factors in the planning or on the time line in a hypothetical sense to evaluate the outcome. The production planner could hypothesize about possible budget cuts, personnel increases or decreases, or less costly ways of making maps. The production planner could test how these events might impact DMA's plans and abilities for meeting customer needs.

The Production Management Segment KBS helps develop plans for the overall production activities which includes the Automatic Feature Extraction of the Data Extraction Segment KBS and the cartographic product creation and finishing services provided by the Product Generation Segment KBS.

LESSONS LEARNED

The following considerations are provided as some of the lessons learned in building the KBSs in the DPS:

- Don't use a KBS when conventional software will satisfy the requirements. KBSs can be very useful in the right applications. One of the factors in the success of a KBS is selecting the right application.

- The nature of KBSs does not exempt them from following sound software life cycle practices. Organization, documentation, and constant review throughout the design and development process of the large scale KBSs are critical to their successful implementation and future growth. The most important consideration during the KBS design process should be the future maintenance of the KBS. If a developer in the earliest stages of the software life cycle can't tell you how a KBS will be maintained, it will fail as a long term application.

- It is important to be aware of both the benefits and limitations of prototyping. Don't confuse the prototype with the final system. Allow ample and separate time for the development of both. Prototypes are often developed in an Artificial Intelligence environment or using expert system shells. The results are impressive, rapidly achieved, and valuable in refining the requirements for the KBS. In contrast to a prototype, a large scale KBS application requires both overhead software to provide the interface to the rest of the system, and optimization to provide reasonable run times.

SUMMARY

In summary, DMA is using KBSs to support the DPS. This is the largest application of KBS to cartography today. This technology, when reasonably applied, can provide great benefits to cartographic processes without incurring risks.